Excited state calculations using phaseless auxiliary-field quantum Monte Carlo: potential energy curves of low lying C_2 singlet states

Wirawan Purwanto, Shiwei Zhang, and Henry Krakauer
Department of Physics, College of William and Mary, Williamsburg, Virginia 23187-8795, USA
(Dated: November 25, 2008)

We show that the recently developed phaseless auxiliary-field quantum Monte Carlo (AFQMC) method can be used to study excited states, providing an alternative to standard quantum chemistry methods. The phaseless AFQMC approach, whose computational cost scales as M^3-M^4 with system size M, has been shown to be among the most accurate many-body methods in ground state calculations. For excited states, prevention of collapse into the ground state and control of the Fermion sign/phase problem are accomplished by the approximate phaseless constraint with a trial wave function. Using the challenging C_2 molecule as a test case, we calculate the potential energy curves of the ground and two low-lying singlet excited states. The trial wave function is obtained by truncating complete active space wave functions, with no further optimization. The phaseless AFQMC results using a small basis set are in good agreement with exact full configuration interaction calculations, while those using large basis sets are in good agreement with experimental spectroscopic constants.

PACS numbers: 71.15.-m,02.70.Ss,21.60.De,31.15.vn,31.50.Bc,31.50.Df,71.15.Qe

I. INTRODUCTION

The ability to calculate electronic excited states of molecules and extended systems is necessary to predict key phenomena and properties of technologically important systems. Compared to ground states, however, the accurate calculation of excitation energies is significantly more difficult. For molecules, a variety of many-body electronic structure quantum chemistry approaches have been developed, typically using a one-particle basis to represent the many-body wave function. For small molecules with modest basis sets, the full configuration-interaction (FCI) method is exact, but FCI is not practical for realistic calculations, since the computational cost scales exponentially as the system size is increased. For larger systems, approximate coupled cluster (CC) methods[1] are the standard, but these methods also have rather steep computational scaling with system size [*e.g.*, $O(M^7)$ for CCSD(T), CC with single and double excitations and perturbational triplets, where M is the number of basis functions]. For extended systems, less accurate approximations based on density functional theory (DFT) and time-dependent DFT have been developed; GW and Bethe-Salpeter type methods have also been shown to be promising.[2] Correlated quantum chemistry methods have also been embedded in DFT calculations to treat extended systems.[3,4] The most commonly applied quantum Monte Carlo (QMC) method in electronic structure has been diffusion Monte Carlo (DMC),[5] which has also been used to compute excited states.[6,7] Compared to ground states, however, the accuracy of the results may depend on the symmetry[6] and show greater sensitivity[8] to the trial wave function used in the fixed-node approximation to control the sign problem and maintain orthogonality.

The recently developed phaseless auxiliary-field quantum Monte Carlo (AFQMC) method[9,10,11,12] is an orbital-based alternative many-body approach. AFQMC can be expressed with respect to any chosen single-particle basis (*e.g.*, gaussians, planewaves, Wannier, etc.), and it exhibits favorable $O(M^3$-$M^4)$ scaling. For ground states, the new AFQMC method has been applied to close to 100 systems, including first- and second-row molecules,[11,12,13] transition metal oxide molecules,[10] simple solids,[9,14] post-d elements,[15] van der Waals systems,[16] and in molecules in which bonds are being stretched or broken.[17,18] In these calculations we have operated largely in an automated mode, inputting only the DFT or Hartree-Fock (HF) solutions as trial wave functions. The method demonstrated excellent accuracy, consistently able to correct errors in the mean-field trial wave function. In molecules, we have found that the accuracy of the phaseless AFQMC is comparable to CCSD(T) near equilibrium geometry and better when bonds are stretched. AFQMC thus provides new opportunities for the efficient and accurate many-body calculations of ground and excited states in both molecular and extended systems.

The seemingly simple C_2 molecule presents a significant challenge for many-body methods.[19,20] The C_2 molecule is difficult because of the strongly multireference nature of the ground state wave function [in which only $\sim$ 70% of the weight is the restricted Hartree-Fock (RHF) determinant] and the presence of nearby low-lying states. The shortcomings of standard quantum chemistry calculations for C_2 were shown by recent benchmark FCI calculations[19] of the potential energy curves (PECs) of its $^1\Sigma_g^+$ ground state and two low-lying singlet excited states. This benchmark shows that most correlated methods based on a single-determinant reference state wave function $|\Phi_r\rangle$ exhibit large nonparallelity errors (NPE—defined as the difference between the maximum and minimum deviations from FCI along the PEC). Spin-restricted CCSD(T) [referred to as RCCSD(T) hereafter] was found to exhibit a large NPE of $98\,\mathrm{m}E_\mathrm{h}$ due to the poor behavior of RHF in the dissociation limit. Spin-unrestricted UCCSD(T), which is usually less accurate near equilibrium, has an NPE of $34\,\mathrm{m}E_\mathrm{h}$. The excited state PECs are not accurately modeled by any of the commonly used single-reference methods, nor by CI including full quadruple substitutions.[19] Similarly, a recent DMC study[20] found that, even in its ground state at equilibrium geometry, the total energy of C_2 showed a large fixed-node error $\sim 40\,\mathrm{m}E_\mathrm{h}$, if a single-determinant trial wave function is used.

As a new QMC method, the phaseless AFQMC provides an alternative route to the sign problem from fixed-node DMC. The random walks take place in a manifold of Slater determinants, in which fermion antisymmetry is automatically maintained in each walker. Applications have indicated that often this reduces the severity of the sign problem and, as a result, the phaseless approximation has weaker reliance on the trial wave function. It is interesting then to test the method for excited states, where QMC calculations depend more on the trial wave function and our knowledge of it is less. The challenging C_2 molecule, where FCI results are available for the modest-sized basis set 6-31G*, provides an excellent test case.

We first describe the AFQMC methodology for ground and excited states. We then make detailed comparions of our C_2 calculated results with the FCI calculations of Ref. 19. Finally, our calculated PECs and spectroscopic constants with large realistic basis sets are presented and compared with experimental results.

II. METHODOLOGY

In this paper, we focus on the lowest-lying C_2 singlet states: the $X\,^1\Sigma_g^+$ ground state and the $B\,^1\Delta_g^+$ and $B'\,^1\Sigma_g^+$ excited states. Since AFQMC uses a projection method to obtain the excited states, collapse to the ground state must be prevented. The $B\,^1\Delta_g^+$ state belongs to a different irreducible representation of the symmetry group of the Hamiltonian than does the $X\,^1\Sigma_g^+$ ground state, but the $B'\,^1\Sigma_g^+$ excited state belongs to the same irreducible representation as the ground state. Both of these cases are discussed below. We first briefly review the phaseless AFQMC method and then discuss the calculation of the excited states.

A. Ground state

Stochastic ground state quantum Monte Carlo (QMC) methods,[5,9,21,22] which are exact in principle, use projection from any reference many-body wave function $|\Phi_r\rangle$, which has non-zero overlap with the ground state. In practice, however, the Fermionic sign problem[5,9,23,24,25] must be controlled to eliminate exponential growth of the variance. For example, in DMC, a single- or multi-reference trial wave function is used to impose approximate nodal boundary conditions of the many-body wave function in electronic configuration space (a Jastrow factor is also included to reduce the stochastic variance). By contrast, phaseless AFQMC samples the many-body wave function with random walkers $\{|\phi\rangle\}$ in the space of Slater determinants, which are expressed in terms of a chosen single-particle basis. Here we use standard quantum chemistry gaussian basis sets.[26] Each $|\phi\rangle$ has the form of a HF or DFT wave function, with the orbitals varying stochastically in the projection. AFQMC controls the sign problem differently, using the complex overlap of the walker $|\phi\rangle$ with a trial/reference wave function $|\Phi_r\rangle$ which is a determinant or a linear combination of determinants.

The ground state energy is obtained from the mixed estimator

$$E_0 = \frac{\langle\Phi_r|\hat{H}|\Psi_0\rangle}{\langle\Phi_r|\Psi_0\rangle} = \lim_{\beta\to\infty}\frac{\langle\Phi_r|\hat{H}e^{-\beta\hat{H}}|\Phi_r\rangle}{\langle\Phi_r|e^{-\beta\hat{H}}|\Phi_r\rangle}\,, \qquad (1)$$

where $\hat{H}$ is the many-body Hamiltonian and $|\Psi_0\rangle$ is the ground state wave function, which is given by imaginary time (β) projection from $|\Phi_r\rangle$. Using the Hubbard-Stratonovich (HS) transformation,[27,28] an importance sampling transformation[9] then expresses the mixed estimator as a stochastic average over the walkers and their Monte Carlo weights w_ϕ,

$$E_0^{\rm MC} = \frac{\sum_\phi w_\phi E_L[\phi]}{\sum_\phi w_\phi}\,, \qquad (2)$$

where the "local energy" E_L is defined as

$$E_L[\phi] \equiv \frac{\langle\Phi_r|\hat{H}|\phi\rangle}{\langle\Phi_r|\phi\rangle}\,. \qquad (3)$$

The energy computed from Eq. (2) is approximate. In addition to a statistical error which can be accurately estimated and reduced with further sampling, there is a systematic error because of the phaseless constraint with $|\Phi_r\rangle$. In other words, the realization of $e^{-\beta\hat{H}}$ in Eq. (1) is approximate,

$$e^{-\beta\hat{H}} \to \widetilde{e^{-\beta\hat{H}}}\,, \qquad (4)$$

because of the constraint on the random walk paths in the Slater determinant space (or equivalently, in the corresponding auxiliary-field space). This is the only approximation in the calculation. The computed ground state energy is not an upper bound.[9,29]

In previous applications, phaseless AFQMC with a single unrestricted Hartree-Fock (UHF) determinant $|\Phi_r^{\rm UHF}\rangle$ was found to often give better overall and more uniform accuracy than CCSD(T) in mapping PECs.[11,15,16] In some cases, however, such as the BH and N_2 molecules, achieving quantitative accuracy of a few m$E_{\rm h}$ for the entire PEC required multi-determinant $|\Phi_r\rangle$.[17] We also find this to be true for the C_2 molecule, as discussed below. Although a multi-reference $|\Phi_r\rangle$ containing N_D determinants increases the computational cost roughly by a factor of N_D, we have found that this is significantly offset by the gain in statistical and systematic accuracy due to the use of a $|\Phi_r\rangle$ that more closely resembles the ground state.[17]

B. Excited states

Calculating the energy of the lowest excited state belonging to an irreducible representation that is different from that of the ground state (*e.g.*, the $B\,^1\Delta_g^+$ state) is straightforward. In this case, we simply choose $|\tilde{\Phi}_r\rangle$ in Eq. (1) to have the symmetry of the desired excited state. Since the Hamiltonian is invariant under the group of symmetry transformations, this

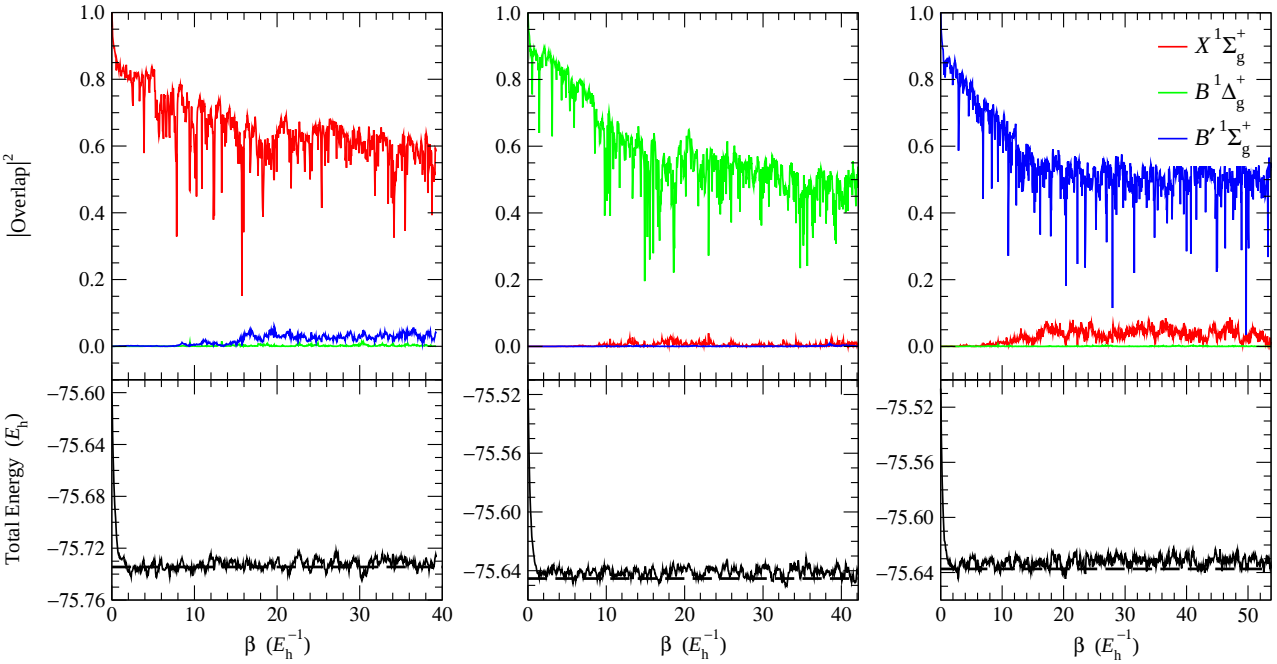

FIG. 1: (Color online) The overlaps of the AFQMC wave function with different states, and the computed total energy, as a function of imaginary time β at interatomic separation 1.25 Å in C_2. The left, center, and right panels show results for the X, B, and B' states, respectively, with AFQMC using the corresponding truncated CASSCF(8,16) wave function as $|\Phi_r\rangle$. In each case, the upper panel shows the three overlap integrals $|\langle\Phi_r^X|\Psi_{\rm MC}\rangle|^2$, $|\langle\Phi_r^B|\Psi_{\rm MC}\rangle|^2$, and $|\langle\Phi_r^{B'}|\Psi_{\rm MC}\rangle|^2$. The lower panels show the computed AFQMC energy, given by Eq. (1), together with the exact FCI energy[19] for comparison (indicated by the horizontal dashed lines). The $|\Phi_r\rangle$ are given by truncated CASSCF(8,16) wave functions. The 6-31G* basis set is used.

projects out, in Eq. (3), any component in the walker determinant $|\phi\rangle$ belonging to a different representation. Although it is usually not possible for a single-reference $|\Phi_r\rangle$ to satisfy the symmetry requirement, multi-reference wave functions can, at least approximately. We use truncated complete active space self-consistent field (CASSCF)[30] wave functions in this study, and our tests indicate that symmetry breaking due to the truncation is small (see below).

For calculations of excited states belonging to the same irreducible representation as the ground state, *e.g.*, the $B'\,^1\Sigma_g^+$ state, we rely on the fact that the corresponding reference wave function is approximately orthogonal to the exact ground state, $\langle\Phi_r^{B'}|\Psi_0^X\rangle \approx 0$. Obtaining accurate AFQMC results for excited states thus depends on using sufficiently accurate excited state trial wave functions. Our results indicate that a multi-reference $|\Phi_r\rangle$ with a modest number of determinants directly taken from a CASSCF calculation is adequate.

Figure 1 illustrates our approach for the first three singlet states in C_2 at a bond length near that of the ground state equilibrium. In AFQMC the wave function is given by

$$|\Psi_{\rm MC}\rangle = \widetilde{e^{-\beta\hat{H}}}|\Phi_r\rangle \sim \sum_\phi w_\phi \frac{|\phi\rangle}{\langle\Phi_r|\phi\rangle}\,, \qquad (5)$$

where $|\Phi_r\rangle$ is the reference function used in the calculation for the phaseless constraint and for importance sampling, and the sum is over the random walker population, $\{w_\phi, |\phi\rangle\}$, at each time slice. From Eq. (5), we can obtain estimates of the overlap integrals $|\langle\Phi_r^s|\Psi_{\rm MC}\rangle|^2$ (where $s = X$, B, or B') to probe the composition of the AFQMC wave function. We normalize $|\Psi_{\rm MC}\rangle$ by explicitly evaluating $\sqrt{\langle\Psi_{\rm MC}|\Psi_{\rm MC}\rangle}$ at each β. Because this involves "undoing" the importance sampling[31] [division by the factor $\langle\Phi_r|\phi\rangle$ on the right-hand side of Eq. (5)], there are large statistical fluctuations, as can be seen in the upper panels. An average population of 1000 walkers is used in these calculations.

The left panel in Fig. 1 shows the ground state calculation, and its upper panel shows a large ($\sim 60\%$) $|\langle\Phi_r^X|\Psi_{\rm MC}\rangle|^2$ overlap of the AFQMC walker population with the reference wave function, as expected. By contrast, the overlap $|\langle\Phi_r^B|\Psi_{\rm MC}\rangle|^2$ is essentially zero, since the symmetry of the $|\Phi_r^B\rangle$ state is different from that of the $|\Phi_r^X\rangle$ ground state, and since symmetry breaking due to truncation of the full CASSCF wave funtion is evidently weak. Moreover, the somewhat larger $|\langle\Phi_r^{B'}|\Psi_{\rm MC}\rangle|^2$ overlap ($\sim 2-5\%$) is not surprising, since $|\Phi_r^{B'}\rangle$ has the same $^1\Sigma_g^+$ symmetry as the ground state. The same trends are observed in the center and right panels of Fig. 1. In the center panel, where the total energy of the B state is calculated, the overlaps of the walker population with the different symmetry X and B' states is extremely small. In each of the three panels, the walker population overlap with the corresponding trial reference state is large. Finally, it is interesting to note that the characteristic imaginary time ($\sim 20-40\,E_h^{-1}$) to reach the asymtoptic value of the overlaps is much larger than the equilibration times ($\sim 3\,E_h^{-1}$) for the total energy. (Section III A presents detailed comparisons with the FCI results.)

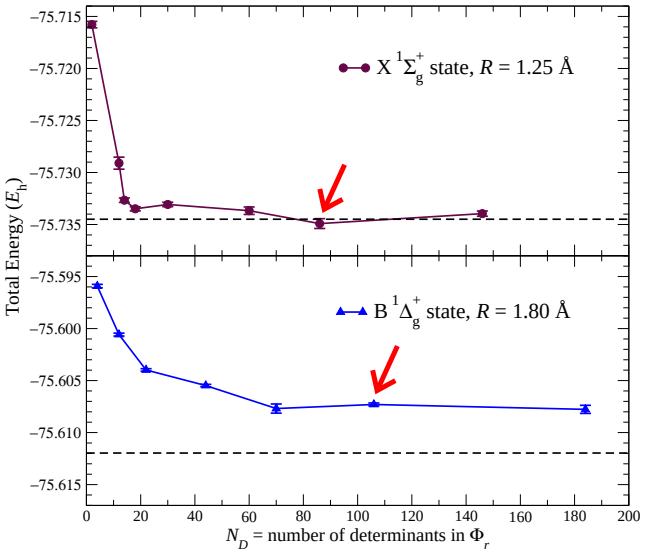

FIG. 2: (Color online) Convergence of the AFQMC energy as a function of the number of determinants N_D in the multi-reference $|\Phi_r\rangle$, obtained by truncating a CASSCF wave function (see text). The upper panel shows the convergence for the X state near its equilibrium at $R = 1.25$ Å. The lower panel corresponds to the B state at $R = 1.8$ Å, a slightly larger bond length than the X-B crossover point at $R \sim 1.7$ Å). The arrows indicate the N_D given by our truncation criterion of 97% integrated weight. The 6-31G* basis set was used. Exact FCI energy[19] is shown as horizontal dashed lines.

C. Computational details

Most of our AFQMC calculations use as $|\Phi_r\rangle$ a truncated CASSCF(8,16) wave function, obtained from the GAMESS quantum chemistry program.[32] The CASSCF wave function is truncated such that the weight (squared coefficient) of the retained determinants is $\sim 97\%$ of the total. Figure 2 plots the computed AFQMC energy as a function of the number N_D of retained determinants (ordered by decreasing weight) for the X and B states, calculated at $R = 1.25$ and $R = 1.8$ Å, respectively. Convergence to the 1 or 2 mE_h level is relatively quick, with the more correlated B state showing somewhat slower convergence. Our truncation criterion above gives $N_D = 86$ and $N_D = 106$ determinants for the X and B states, respectively. The corresponding AFQMC energies are seen to be well converged with respect to the truncation.[33]

Our AFQMC calculations use the local energy formalism,[9,31] using standard gaussian-type basis sets.[26] Reference wave functions were obtained using GAMESS,[32] and the one- and two-body matrix elements were obtained using a modified NWCHEM code.[34] A mean-field background subtraction is applied to the Hamiltonian prior to the HS transformation, which improves the computational efficiency and reduces systematic errors.[11,35] In most of our AFQMC calculations, we use $\Delta\tau = 0.01\,E_h^{-1}$. We confirmed that the resulting Trotter error is less than 1 mE_h by calculations at multiple $\Delta\tau$ values at selected geometries in both the 6-31G* and larger basis sets.

All runs use an average population of about 100 random walkers, with initial population generated from the RHF wave function or "broken symmetry" RHF.[19] (A short phaseless AFQMC projection is first invoked for $\beta \sim 1\,E_h^{-1}$, with the RHF wave function as $|\Phi_r\rangle$ and its copies to form the initial population; the resulting population, which is purely spin-singlet,[18] is then fed into the regular calculation with the CASSCF $|\Phi_r\rangle$.) Typical runs have an equilibration phase of $\beta \sim 10\,E_h^{-1}$ and then a growth phase of $\beta \sim 10\,E_h^{-1}$, in which the trial energy is adjusted via the growth estimator[36] and set for the rest of the simulation. A measurement of $\beta \sim 150\,E_h^{-1}$ is needed to achieve a statistical accuracy of 1 mE_h. To give an idea of the present computational cost, such a calculation with the cc-pVTZ basis at a single geometry (with $N_D \sim 100$ in the trial wave function) requires approximately 2.5 days on one core of a 2.2GHz Opteron processor. The current implementation simply imports[11] one- and two-body gaussian matrix elements from other quantum chemistry programs. The Hamiltonian and overlap integrals are treated as dense matrices. No special properties of the underlying gaussian basis set are exploited.

Our focus with the phaseless AFQMC has so far been on establishing the basic framework in a variety of systems, and testing its systematic accuracy and robustness. In this paper our main purpose is to present a proof of concept for excited state calculations. Although the present implementation (with gaussian basis sets and a density decomposition of the two-body interaction) has shown excellent accuracy, there remains considerable flexibility in the choice of the one-particle basis and the form of the HS transformation. It is possible that exploiting the flexibility can lead to significant further improvement in accuracy and computational efficiency.

III. RESULTS AND DISCUSSION

As bonds are stretched in the dissociation of molecules, accurate treatment of strong electronic correlations becomes important, especially in the intermediate regime when bonds first begin to break. The C_2 molecule is a particularly challenging example, and this trend is seen clearly below. The ground state at equilibrium geometry is already nontrivial,[20] but as the bond is stretched in the ground state and across all geometries in the excited states, the systematic errors grow significantly in all calculations. We first compare our phaseless AFQMC results with benchmark FCI calculations.[19] (More detailed discussion about recent theoretical calculations on C_2 molecule can be found in Ref. 37.) Calculated PECs and spectroscopic constants from realistic basis sets are then presented and compared with experimental results.

A. Comparison with benchmark FCI results

Figure 3 compares the phaseless AFQMC ground state PEC with the FCI results from Ref. 19, using the 6-31G* gaussian basis set. In order to benchmark our AFQMC calculations, which do not employ the frozen core approximation, we estimate a frozen-core correction to FCI using the difference

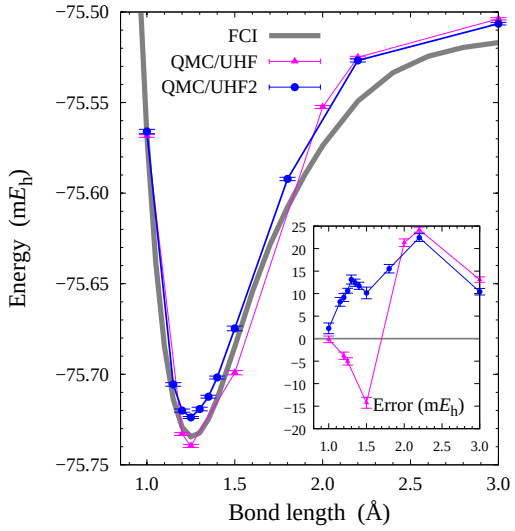

FIG. 3: (Color online) C_2 ground state PEC: comparison of phaseless AFQMC using single-determinant UHF and two-determinant UHF2 reference wave functions (see text) with benchmark FCI results from Ref. 19. All calculations used the 6-31G* gaussian basis set.

of UCCSD(T) energies with and without the frozen-core approximation. The AFQMC/UHF PEC was calculated using a single determinant $|\Phi_r\rangle$ from a UHF calculation for the singlet ground state. The AFQMC/UHF2 PEC used the simplest multi-reference $|\Phi_r\rangle$ consisting of two determinants,

$$|\mathrm{UHF2}\rangle \propto |\mathrm{UHF}\rangle + \alpha|\mathrm{UHF}_{ax}\rangle\,, \tag{6}$$

where the parameter α is variationally optimized. In Eq. (6), $|\mathrm{UHF}\rangle$ is the usual UHF wave function, which is a broken symmetry state with opposite spins on the two C atoms, but which preserves the axial symmetry of the molecule. The $|\mathrm{UHF}_{ax}\rangle$ state[38] breaks the axial symmetry of the molecule and is analogous to the "antiferromagnetic solution" found in a local density approximation calculation in Ref. 39. Near equilibrium, the energy of $|\mathrm{UHF}\rangle$ is lower than that of $|\mathrm{UHF}_{ax}\rangle$, but, as the bond is stretched beyond $R \sim 1.5$ Å, the $|\mathrm{UHF}_{ax}\rangle$ energy becomes lower. The combined $|\mathrm{UHF2}\rangle$ state in Eq. (6) has a variational energy that is $\sim 10\,\mathrm{m}E_\mathrm{h}$ lower than the $|\mathrm{UHF}\rangle$ near equilibrium and $\sim 20\,\mathrm{m}E_\mathrm{h}$ lower at $R \sim 1.8$ Å.

The AFQMC/UHF PEC has an NPE of $38\,\mathrm{m}E_\mathrm{h}$ while AFQMC/UHF2 has a smaller NPE of $20\,\mathrm{m}E_\mathrm{h}$. The NPEs of standard quantum chemistry calculations are shown in Table I, together with those of AFQMC. The single reference AFQMC/UHF NPE is thus seen to be significantly better than RCCSD(T) and comparable to UCCSD(T). Using only two determinants, AFQMC/UHF2 has a smaller NPE than the CISDTQ result. (AFQMC/UHF2 has a lower energy at large bond lengths due to the change in the leading terms of the FCI wave function from the single RHF-like configuration near equilibrium to a two-determinant configuration in the dissociation limit.) Nevertheless, an NPE of $\sim 20\,\mathrm{m}E_\mathrm{h}$ is unacceptably large for a high-level method such as QMC. We also note

TABLE I: Nonparallelity error (NPE) of standard quantum chemistry methods for the C_2 ground state PEC (taken from Ref. 19), compared with that of the phaseless AFQMC method using UHF, two-determinant UHF2, and truncated CASSCF(8,16) trial wave functions. The range of bond lengths is $R = 0.9 - 3.0$ Å. All calculations used the 6-31G* gaussian basis set.

Method	NPE (mE_h)
RHF	212
UHF	78
MP2	130
RCCSD(T)	98
UCCSD(T)	34
CISD	116
CISDT	51
CISDTQ	26
AFQMC/UHF	38
AFQMC/UHF2	20
AFQMC/CASSCF	7

that removing spin-contamination in the walker population, as discussed in Ref. 18, does not yield significant improvements to either AFQMC/UHF or AFQMC/UHF2. This indicates that these $|\Phi_r\rangle$ are themselves poor.

We now show that the phaseless AFQMC results become much more accurate when multi-reference $|\Phi_r\rangle$ are used. As described in Sec. II C, these are truncated CASSCF(8,16) wave functions, with no further optimization. Figure 4 compares the AFQMC/CASSCF and FCI calculated PECs. We see that all three PECs are mapped out accurately, including the B' PEC which is of the same symmetry as the ground state X. The overall accuracy of the AFQMC PECs for both the ground state and the excited states is better than $8\,\mathrm{m}E_\mathrm{h}$ for all but one point, the smallest bond length in B. The X and B crossover at $R \sim 1.7$ Å and the B and B' crossover at $R \sim 1.1$ Å are both accurately described. With the truncated CASSCF trial wave function, the energies also appear to be variational in all cases, while the AFQMC/UHF energies are not.

Table II presents the spectroscopic constants corresponding to the PECs in Fig. 4. Since CCSD(T) is poor in the dissociation limit, the dissociation energies D_e in Table II are calculated using the energy of the free C atom for each method. For AFQMC, this somewhat improves the comparison of D_e with FCI, since the AFQMC energy of the C atom is more accurate than that from the value in Fig. 4 in the dissociation limit, where the truncated CASSCF(8,16) $|\Phi_r\rangle$ is not very good. The spectroscopic constants obtained from energy expectation values of the truncated CASSCF $|\Phi_r\rangle$ have substantial fitting uncertainties, which originate from noise in the determinant truncation. Both the AFQMC/CASSCF and the full CASSCF results are in very good agreement with the exact FCI results, while those based on energy expectation values of the truncated CASSCF $|\Phi_r\rangle$ are significantly worse. This shows that AFQMC improves significantly over the truncated CASSCF trial wave functions.

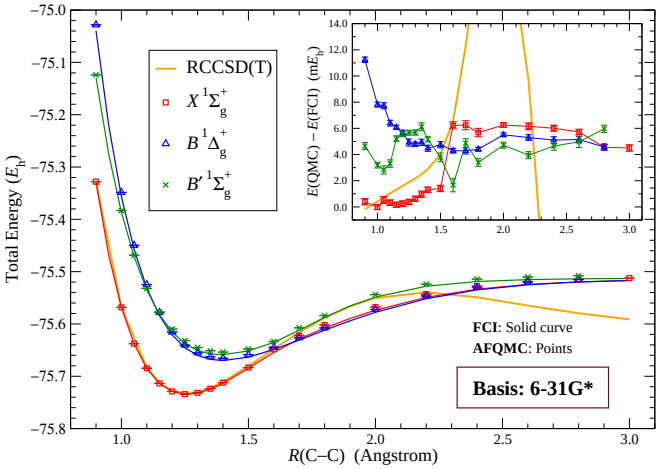

FIG. 4: (Color online) AFQMC/CASSCF PECs for the three lowest lying singlet states in C_2, using multi-reference truncated CASSCF(8,16) $|\Phi_r\rangle$, compared with FCI results.[19] FCI results are given by the solid curves and AFQMC results are given by symbols with error bars. RCCSD(T) results for the ground state are also shown as the orange line. The inset shows deviations from FCI in mE_h. All calculations used the 6-31G* gaussian basis set.

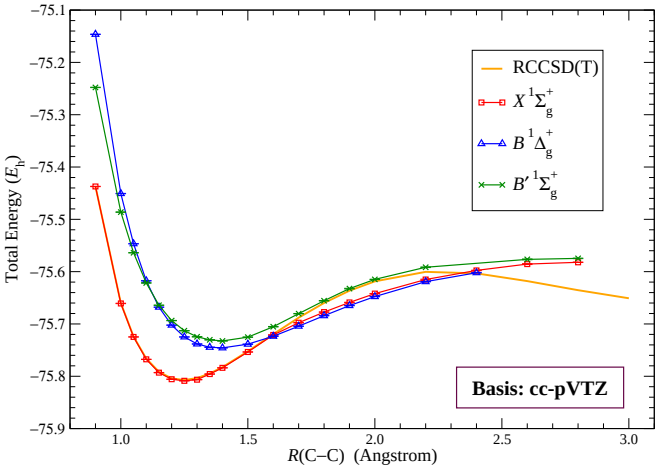

FIG. 5: (Color online) C_2 AFQMC/CASSCF PECs using $|\Phi_r\rangle$ given by a truncated CASSCF(8,16) wave function. For comparison, the RCCSD(T) ground state PEC is also shown. All calculations used the cc-pVTZ basis.[26]

TABLE II: AFQMC/CASSCF calculated C_2 spectroscopic constants, corresponding to Fig. 4 using multi-reference truncated CASSCF(8,16) $|\Phi_r\rangle$. For comparison, results from FCI,[19] CCSD(T), the *full* CASSCF, and from the energy expectation values of the *truncated* CASSCF $|\Phi_r\rangle$ are also shown. All calculations used the 6-31G* basis set. Results for the equilibrium bond length r_e (in Å), vibrational frequency ω_e (in cm^{-1}), and ground state dissociation energy D_e (in eV) are shown. For excited states, the excitation energy T_e is defined as the energy difference between the minima of the excited and ground states. Combined statistical and fitting errors in AFQMC are shown in parantheses, while pure fitting uncertainties are shown in square brackets.

	CASSCF(8,16)				
	full	truncated	CCSD(T)	QMC	FCI
$X\,^1\Sigma_g^+$ ground state					
r_e	1.2563	1.271[4]	1.2577	1.2567(3)	1.2581
ω_e	1893	1705[76]	1869	1888(12)	1863
D_e	6.530	5.12[1]	5.953	6.085(5)	6.030
$B\,^1\Delta_g^+$ excited state					
r_e	1.3997	1.39[3]		1.4004(9)	1.4000
ω_e	1414[1]	1360[332]		1394(13)	1391
T_e	1.712	1.75[8]		1.874(6)	1.761
$B'\,^1\Sigma_g^+$ excited state					
r_e	1.3868	1.399[7]		1.4009(12)	1.3931
ω_e	1463	1453[127]		1325(32)	1394
T_e	1.920	1.91[3]		2.189(6)	2.058

B. Realistic C_2 results using large basis sets

In this subsection, we present phaseless AFQMC/CASSCF PECs with large basis sets and compare with experimental spectroscopic values. Figure 5 shows AFQMC/CASSCF PECs for the three lowest lying singlet states, using truncated CASSCF(8,16) wave function as $|\Phi_r\rangle$, in the cc-pVTZ basis set.[26] For comparison, the figure also shows the RCCSD(T) ground state PEC, which was calculated using the NWCHEM[34] computer program. Table III presents the corresponding spectroscopic constants and experimental values. We have also computed the ground state spectroscopic constants using the larger cc-pVQZ basis set,[26] shown in Table IV. As seen from these two tables, the residual finite basis set error is small and the computed spectroscopic constants are nearly converged. (Ground-state results obtained using the cc-pV5Z basis set by the contracted multi-reference CI (CMRCI) method[40] are also included in Table IV.)

The calculated AFQMC results in Tables III and IV are in very good agreement with experimental values. The CASSCF(8,16) results, i.e., using the *full* CASSCF WF are also quite good. Each CASSCF(8,16) WF consists of 414864 determinants in the D_{2h} space group. The truncated CASSCF $|\Phi_r\rangle$ used in AFQMC, however, uses only the first $52-275$ determinants, which amounts to 97% of the total CASSCF determinant weight. At smaller distances, the correlation effects are small and thus fewer determinants are needed. At large bond lengths, there are more determinants included in the $|\Phi_r\rangle$, indicative of stronger correlation effects. As seen in Table III, AFQMC significantly improves the results when compared to those obtained from the variational estimate using the *truncated* CASSCF WF. The level of agreement here between the AFQMC results with larger basis sets and experiment is consistent with that in the smaller basis set between AFQMC and FCI in the previous subsection.

IV. SUMMARY

We have shown that molecular excited-state calculations are possible with the phaseless auxiliary-field QMC method. Using CASSCF trial WFs, the method delivers accurate PECs in the challenging C_2. The computed spectroscopic constants

TABLE III: AFQMC/CASSCF calculated C_2 spectroscopic constants compared to experiment. Results from CCSD(T), the *full* CASSCF, and from the energy expectation values of the *truncated* CASSCF wave functions are also shown. Conventions are as in Table II. All calculations used the cc-pVTZ basis set.

	CASSCF(8,16)				
	full	truncated	CCSD(T)	QMC	Expt.
$X\,^1\Sigma_g^+$ ground state					
r_e	1.2479	1.262[3]	1.2508	1.2462(9)	1.2425
ω_e	1862	1766[67]	1842	1884(17)	1855
D_e	6.53	4.64[1]	6.03	6.32(1)	6.33
$B\,^1\Delta_g^+$ excited state					
r_e	1.397[1]	1.416[7]		1.391(1)	1.3855
ω_e	1395[23]	1351[101]		1376(23)	1407
T_e	1.511[5]	1.38[3]		1.723(7)	1.498
$B'\,^1\Sigma_g^+$ excited state					
r_e	1.381	1.398[4]		1.393(1)	1.3774
ω_e	1489[4]	1392[63]		1441(12)	1424
T_e	1.779[2]	1.68[2]		2.082(8)	1.910

TABLE IV: AFQMC/CASSCF calculated C_2 ground state spectroscopic constants compared to experiment. Conventions are as in Table II. Calculations used the cc-pVQZ basis set (except CMRCI which used the cc-pV5Z basis).

	CASSCF(8,16)					
	full	truncated	CCSD(T)	CMRCI[a]	QMC	Expt.
$X\,^1\Sigma_g^+$ ground state						
r_e	1.2452	1.262[3]	1.2459	1.2467	1.244(1)	1.2425
ω_e	1868	1759[29]	1852	1853	1850(21)	1855
D_e	6.57	4.69[1]	6.19	6.29	6.41(1)	6.33

[a] Values from analytical fitting in Ref. 40

for the lowest three singlet states, two of which have the same symmetry, are in very good agreement with experiment.

Acknowledgments

The work was supported in part by DOE (DE-FG05-08OR23340 and DE-FG02-07ER46366). H.K. also acknowlesges support by ONR (N000140510055 and N000140811235), and W.P. and S.Z. by NSF (DMR-0535592). Calculations were performed at the Center for Piezoelectrics by Design, and the College of William & Mary's SciClone cluster. We are grateful to Wissam Al-Saidi and Eric Walter for many useful discussions.

[1] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).
[2] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[3] T. Klüner, N. Govind, Y. A. Wang, and E. A. Carter, J. Chem. Phys. **116**, 42 (2002).
[4] Q.-M. Hu, K. Reuter, and M. Scheffler, Phys. Rev. Lett. **98**, 176103 (2007).
[5] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001), also see the references therein.
[6] W. M. C. Foulkes, R. Q. Hood, and R. J. Needs, Phys. Rev. B **60**, 4558 (1999).
[7] J. C. Grossman, M. Rohlfing, L. Mitas, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **86**, 472 (2001).
[8] F. Schautz, F. Buda, and C. Filippi, J. Chem. Phys. **121**, 5836 (2004).
[9] S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003).
[10] W. A. Al-Saidi, H. Krakauer, and S. Zhang, Phys. Rev. B **73**, 075103 (2006).
[11] W. A. Al-Saidi, S. Zhang, and H. Krakauer, J. Chem. Phys. **124**, 224101 (2006).
[12] M. Suewattana, W. Purwanto, S. Zhang, H. Krakauer, and E. J. Walter, Phys. Rev. B **75**, 245123 (2007).
[13] S. Zhang, H. Krakauer, W. A. Al-Saidi, and M. Suewattana, Comput. Phys. Commun. **169**, 394 (2005).
[14] H. Kwee, S. Zhang, and H. Krakauer, Phys. Rev. Lett. **100**, 126404 (2008).
[15] W. A. Al-Saidi, H. Krakauer, and S. Zhang, J. Chem. Phys. **125**, 154110 (2006).
[16] W. A. Al-Saidi, H. Krakauer, and S. Zhang, J. Chem. Phys. **126**, 194105 (2007).
[17] W. A. Al-Saidi, S. Zhang, and H. Krakauer, J. Chem. Phys. **127**, 144101 (2007).
[18] W. Purwanto, W. A. Al-Saidi, H. Krakauer, and S. Zhang, J. Chem. Phys. **128**, 114309 (2008).
[19] M. L. Abrams and C. D. Sherrill, J. Chem. Phys. **121**, 9211 (2004).
[20] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007).
[21] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
[22] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, J. Chem. Phys. **77**, 5593 (1982).
[23] D. M. Ceperley and B. J. Alder, J. Chem. Phys. **81**, 5833 (1984).
[24] S. Zhang and M. H. Kalos, Phys. Rev. Lett. **67**, 3074 (1991).
[25] S. Zhang, in *Quantum Monte Carlo Methods in Physics and Chemistry*, edited by M. P. Nightingale and C. J. Umrigar (Kluwer Academic Publishers, 1999), cond-mat/9909090.

[26] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, J. Chem. Inf. Model. **47**, 1045 (2007).
[27] J. Hubbard, Phys. Rev. Lett. **3**, 77 (1959).
[28] R. D. Stratonovich, Dokl, Akad. Nauk. SSSR **115**, 1907 (1957).
[29] J. Carlson, J. E. Gubernatis, G. Ortiz, and S. Zhang, Phys. Rev. B **59**, 12788 (1999).
[30] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, Chem. Phys. **48**, 157 (1980).
[31] W. Purwanto and S. Zhang, Phys. Rev. E **70**, 056702 (2004).
[32] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, et al., J. Comput. Chem. **14**, 1347 (1993).
[33] The CASSCF natural orbitals were also used to construct an alternative determinantal expansion of the full CASSCF wave function, which then was truncated in the usual way. For large $N_D \gtrsim 60$, truncating the alternative CAS wave function yielded essentially identical AFQMC energies as the original truncation. For smaller N_D, however, convergence of the AFQMC energy was slower than with the original method of truncation.
[34] E. Aprà, T. Windus, T. Straatsma, E. Bylaska, W. de Jong, S. Hirata, M. Valiev, M. Hackler, L. Pollack, K. Kowalski, et al., *NWChem, A Computational Chemistry Package for Parallel Computers, Version 4.6*, Pacific Northwest National Laboratory, Richland, Washington 99352-0999, USA (2004).
[35] W. Purwanto and S. Zhang, Phys. Rev. A **72**, 053610 (2005).
[36] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. B **55**, 7464 (1997).
[37] C. D. Sherrill and P. Piecuch, J. Chem. Phys. **122**, 124104 (2005).
[38] The spin-up valence orbitals of $|\mathrm{UHF}_{ax}\rangle$ can be described as $2(s-p_z)\sigma_g$, $[2(s+p_z)\sigma_u^* + 2p_x\pi_g^*]$, $(2p_z\sigma_g + 2p_x\pi_u)$, $2p_y\pi_u$.
[39] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).
[40] M. Boggio-Pasqua, A. Voronin, P. Halvick, and J.-C. Rayez, J. Mol. Struct. (THEOCHEM) **531**, 159 (2000).